Available at WWW.MATHEMATICSWEB.ORG POWERED BY SCIENCE DIRECT

ELSEVIER

Journal of Computational and Applied Mathematics 156 (2003) 433–456

JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS

www.elsevier.com/locate/cam

Extension and further development of the differential calculus for matrix norms with applications

L. Kohaupt

Prager Strasse 9, D-10779 Berlin, Germany

Received 5 October 2001; received in revised form 24 June 2002

Abstract

In this paper, the differential calculus for the operator norms $\|\cdot\|_p$, $p\in\{1,2,\infty\}$, of the fundamental matrix or evolution $\Phi(t)=e^{At}$, $t\geqslant 0$, of a complex $n\times n$ matrix A, introduced by the author in a former paper, is extended to m times continuously differentiable matrix functions $\chi(t)$, $t\geqslant 0$, and developed further for other p-norms $|\cdot|_p$, $1<p<\infty$. Results similar to those for $\Phi(t)$ are obtained. In addition, for this function $\Phi(t)$, formulae for the first two logarithmic derivatives $D_+^1|\Phi(0)|_p$ and $D_+^2|\Phi(0)|_p$, $1<p<\infty$, are obtained as special cases. Also, upper bounds on the discrete evolution $\Psi(t)$, $t\geqslant 0$ (that is, a matrix power function) and on the difference (or remainder) $R(t)=\Phi(t)-\Psi(t)$, $t\geqslant 0$, are derived. The discrete evolution occurs when a step-by-step method is employed to approximate the exact solution of the initial-value problem $\dot{x}(t)=Ax(t)$, $x(0)=x_0$, which here models a vibration problem. The results are applied to the computation of the optimal upper bounds on $\|R(t)\|_\infty$, $\|R(t)\|_2$, and $|R(t)|_2$.

Keywords: Differential calculus for norms of a matrix function; Remainder; Best upper bound; Logarithmic derivative; Application to vibration problem

1. Introduction

The solution of the initial-value problem $\dot{x}(t)=Ax(t)$, $x(0)=x_0$, is given by $x(t)=\Phi(t)x_0$ where $\Phi(t)=e^{At}$ is the fundamental matrix or evolution of the $n\times n$ matrix A. One classical upper bound on $\|\Phi(t)\|$ is known as $\|\Phi(t)\|\leqslant M_\varepsilon e^{(v(A)+\varepsilon)t}$, $t\geqslant 0$, where $v(A)$ means the spectral abscissa of matrix A. Here, however, the constant M_ε obtained by classical methods is not optimal. The minimal M_ε can be computed by the differential calculus for norms developed by the author in [12] for operator norms $\|\cdot\|_p$ with $p\in\{1,2,\infty\}$.

E-mail address: kohaupt@tfh-berlin.de (L. Kohaupt).

0377-0427/03/$ - see front matter
doi:10.1016/S0377-0427(03)00385-6

The *aim* of this paper is twofold, namely:

(a1) to *extend* the differential calculus for the operator p-norms $\|\cdot\|_p$, $p\in\{1,2,\infty\}$, from $\Phi(t)$ to general matrix functions $\chi(t)$, and
(a2) to *develop further* a corresponding differential calculus for other p-norms $|\cdot|_p$, $1<p<\infty$.

One is led to the *first aim* (a1) by considering two additional matrix functions, namely

1. the discrete evolution $\Psi(t)$, $t\geqslant 0$ (which appears when the above initial-value problem is discretized, e.g., by a finite-difference method), and
2. the difference (or remainder) $R(t):=\Phi(t)-\Psi(t)$, $t\geqslant 0$.

For both matrix functions, upper bounds on their norms are derived. To compute the corresponding minimal constants M_ε, the differential calculus for the norms $\|\cdot\|_p$, $p\in\{1,2,\infty\}$, has to be carried over from the matrix function $\Phi(t)$ to the functions $\Psi(t)$ and $R(t)$. In order to be more general, the differential calculus is *extended* here to matrix functions $\chi(t)$ that are sufficiently differentiable. For $p\in\{1,\infty\}$, this extension is straightforward; for $p=2$, more effort is necessary because we have to consider the case $\chi(t_0)=0$, which happens, e.g., for $\chi=R=\Phi-\Psi$, whereas formerly we had always $\Phi(t_0)\neq 0$.

To the *second aim* (a2), namely to *develop further* the differential calculus for other matrix p-norms $|\cdot|_p$, $1<p<\infty$, one is led in a natural way because only for $p\in\{1,2,\infty\}$ formulae for the norms $\|A\|_p$, $A\in\mathbb{C}^{n\times n}$, are known. From the obtained new formulae for $D_+^1|\chi(t)|_p$ and $D_+^2|\chi(t)|_p$, also the first two logarithmic derivatives $D_+^1|\Phi(0)|_p$ and $D_+^2|\Phi(0)|_p$, $1<p<\infty$, follow as special cases.

The results are applied to a vibration problem and are illustrated by graphics and numerical values.

More precisely, the paper is structured as follows. For the norms $\|\cdot\|_p$, $p\in\{1,2,\infty\}$ and $|\cdot|_p$, $1<p<\infty$, of matrix functions $\chi(t)$, in Section 2, local regularity properties are stated, and in Section 3, formulae for right derivatives are obtained. In Section 4, upper bounds on $\|\Psi(t)\|$ and on $\|R(t)\|$ are determined by classical methods. Section 5 is the application part. We consider again the vibration problem of Ref. [12] and restrict ourselves essentially to the function $R(t)$ because the shape of the upper bound on $\|R(t)\|$ is very different from that on $\|\Phi(t)\|$. (The figures for the upper bounds on $\|\Psi(t)\|$ are omitted since they strongly resemble those on $\|\Phi(t)\|$.) We use the differential calculus for norms of matrix functions, derived in the earlier sections, to obtain the optimal upper bounds on $\|R(t)\|_\infty$, $\|R(t)\|_2$, and $|R(t)|_2$. The Refs. [2,3,6–9,14–21,23] are given even though they are not directly used in this paper in order to provide the reader with some additional material helpful in the present subject.

2. Local regularity of norms of matrix functions

In [12], for $t\mapsto\Phi(t)$, $t\geqslant 0$, we have shown—loosely speaking—that for every $t_0\geqslant 0$ and for $p\in\{\infty,2\}$ the function $t\mapsto\|\Phi(t)\|_p$ is real analytic in some neighbourhood $[t_0,t_0+\Delta t_0]$. The case $p=1$ can easily be reduced to the case $p=\infty$ by interchanging the column index and the row index. Corresponding results hold when Φ is replaced by any analytic function $\chi\in\mathbb{C}^{n\times n}$ or more generally when $\chi\in\mathbb{C}^{n\times n}$ is sufficiently often continuously differentiable.

In this section, differentiability results for general matrix functions $\chi(t)$ are derived *extending* the results of Ref. [12] for the norms $\|\cdot\|_p,\ p\in\{1,2,\infty\}$, and *developing further* the results for other p-norms $|\cdot|_p,\ 1<p<\infty$.

Thus, we have

Lemma 1 ($\|\cdot\|_p,\ p\in\{1,2,\infty\}$, complex matrix function). *Let $m\in\mathbb{N},\ t_0\in\mathbb{R}_0^+$ and $\chi:\mathbb{R}_0^+\mapsto\mathbb{C}^{n\times n}$ be a matrix function that is m times continuously differentiable. Further, suppose* ***additionally*** *that for $p\in\{1,\infty\}$ each two components of $\chi(t)$ and for $p=2$ each two eigenvalues of $\chi^*(t)\,\chi(t)$ be either identical or intersect each other at most finitely often near t_0. Then, there exists a number $\Delta t_0>0$ and a function $t\mapsto\hat{\chi}(t)$, which is real and m times continuously differentiable on $[t_0,t_0+\Delta t_0]$, such that $\hat{\chi}(t)=\|\chi(t)\|_p$ for all $t\in[t_0,t_0+\Delta t_0]$.*

Proof. The proof is similar to that of [12, Lemma 1] and [13, Lemma 3]. In Lemma 1 and all other cases of the norms $\|\cdot\|_p$, we have to make the additional hypothesis for all $p\in\{1,2,\infty\}$ since the maximum on n numbers has to be formed in all these cases. □

Supplement 2. ($\|\cdot\|_p,\ p\in\{1,2,\infty\}$, *complex matrix function*). *If χ is analytic for $t\geqslant 0$ (or in a neighbourhood of the considered point $t_0\in\mathbb{R}_0^+$), then the additional hypothesis in Lemma 1 can be dropped.*

Proof. The proof is similar to the proof of [13, Supplement 4]. □

For the matrix operator norms $\|\cdot\|_p$, no explicit representation $\|A\|_p,\ A\in\mathbb{C}^{n\times n}$, is known for $p\notin\{1,2,\infty\}$. So, in addition, apart from considering these matrix operator p-norm we introduce other p-norms $|\cdot|_p,\ 1<p<\infty$, and prove similar results as mentioned above. For the sake of completeness, this is also done in the present paper.

Let $B=(b_{ij})\in\mathbb{C}^{n\times n}$ and

$$|B|_p := \left(\sum_{i,j=1}^{n}|b_{ij}|^p\right)^{1/p},\quad 1<p<\infty. \tag{1}$$

Then, $|\cdot|_p$ are norms on $\mathbb{C}^{n\times n}$ which are not, however, operator norms. Let $u\in\mathbb{C}^n$. Then,

$$\|Bu\|_p\leqslant|B|_p\,\|u\|_q,$$

where q is the number conjugate to p, that is, $1/p+1/q=1$ or $q=p/(p-1)$. For $1<p\leqslant 2$, one has additionally

$$\|u\|_q\leqslant\|u\|_p,$$

and

$$|BC|_p\leqslant|B|_p\,|C|_p,$$

where $B, C \in \mathbb{C}^{n\times n}$ so that $|\cdot|_p$ is a submultiplicative matrix norm if $1 < p \leqslant 2$. For $p=2$, this follows from Cauchy–Schwarz's and for $1 < p < 2$, from Jansen's inequality (cf. [22, p. 6]). Moreover, it is clear that

$$\|B\|_p \leqslant |B|_p, \ 1 < p \leqslant 2.$$

We mention that, in [19], the norm $|\cdot|_2$ is called *Schur norm*, and in [24] it is also called *Euclidian norm*. Other authors call it *Frobenius norm*.

For the norms $|\cdot|_p$, the following lemma holds true.

Lemma 3 ($|\cdot|_p$, $1 < p < \infty$). *Let $1 < p < \infty$, $t_0 \in \mathbb{R}_0^+$, and $m \in \mathbb{N}$. Further, for every $t \geqslant 0$, let $\chi(t) \in \mathbb{C}^{n\times n}$ where $t \mapsto \chi(t)$, $t \geqslant 0$, is m times continuously differentiable.*

Then, there exists a number $\Delta t_0 > 0$ and a function $t \mapsto \hat{\chi}(t)$, which is real and m times continuously differentiable on $[t_0, t_0 + \Delta t_0]$, such that $\hat{\chi}(t) = |\chi(t)|_p$ for all $t \in [t_0, t_0 + \Delta t_0]$.

Proof. We leave the proof to the reader since it is similar to that of [13, Lemma 3]. □

3. Formulae for the right derivatives of norms

In this section, formulae for the right derivatives of general matrix functions $\chi(t)$ are obtained *extending* the results of [12] for the norms $\|\cdot\|_p$, $p \in \{\infty, 2\}$, and *developing further* the results for other p-norms $|\cdot|_p$, $1 < p < \infty$. As a special case, also the first two logarithmic derivatives $D_+^1|\Phi(0)|_p$ and $D_+^2|\Phi(0)|_p$ are derived.

3.1. *Matrix functions $t \mapsto \chi(t)$ in the operator norms $\|\cdot\|_p$, $p \in \{\infty, 2\}$*

One obtains the formulae for the right derivatives of the matrix operator p-norms $\|\cdot\|_p$, $p \in \{\infty, 2\}$, when $A^k\,\Phi(t_0)$ in the formulae of [12] are replaced by the derivatives $D^k\chi(t_0)$, $k = 0, 1, 2, \ldots$. Even though this can easily be done by the reader, we give here the formulae for ease of reference in the future. Since for general matrix functions also the case $D^k\chi(t_0) = 0$ may occur, some additional considerations have to be made in the case $p = 2$.

$p = \infty$: *Complex $n \times n$ matrix $\chi(t)$*. Let $t_0 \in \mathbb{R}_0^+ = \{t \in \mathbb{R} \mid t \geqslant 0\}$, and for $i, j = 1, \ldots, n$ define the functionals

$$\lambda_{ij}^{(0)}[\chi, t_0] := |\chi_{ij}(t_0)|, \tag{2}$$

$$\lambda_{ij}^{(1)}[\chi, t_0] := \begin{cases} \dfrac{\operatorname{Re}\chi_{ij}(t_0)\operatorname{Re}(D\chi)_{ij}(t_0) + \operatorname{Im}\chi_{ij}(t_0)\operatorname{Im}(D\chi)_{ij}(t_0)}{|\chi_{ij}(t_0)|}, & \chi_{ij}(t_0) \neq 0, \\ |(D\chi)_{ij}(t_0)|, & \chi_{ij}(t_0) = 0, \end{cases} \tag{3}$$

$$\lambda_{ij}^{(2)}[\chi, t_0] := \begin{cases} \dfrac{|(D\chi)_{ij}(t_0)|^2 + \operatorname{Re}\chi_{ij}(t_0)\operatorname{Re}(D^2\chi)_{ij}(t_0) + \operatorname{Im}\chi_{ij}(t_0)\operatorname{Im}(D^2\chi)_{ij}(t_0)}{|\chi_{ij}(t_0)|} & \\ \quad - \dfrac{[\operatorname{Re}\chi_{ij}(t_0)\operatorname{Re}(D\chi)_{ij}(t_0) + \operatorname{Im}\chi_{ij}(t_0)\operatorname{Im}(D\chi)_{ij}(t_0)]^2}{|\chi_{ij}(t_0)|^3}, & \chi_{ij}(t_0) \neq 0, \\ \dfrac{\operatorname{Re}(D\chi)_{ij}(t_0)\operatorname{Re}(D^2\chi)_{ij}(t_0) + \operatorname{Im}(D\chi)_{ij}(t_0)\operatorname{Im}(D^2\chi)_{ij}(t_0)}{|(D\chi)_{ij}(t_0)|}, & \chi_{ij}(t_0) = 0,\ (D\chi)_{ij}(t_0) \neq 0, \\ |(D^2\chi)_{ij}(t_0)|, & \chi_{ij}(t_0) = 0,\ (D\chi)_{ij}(t_0) = 0, \end{cases} \tag{4}$$

where $(D\chi)_{ij}(t_0) := [D\chi(t_0)]_{ij}$, and so on. Let

$$\lambda_i^{(k)}[\chi, t_0] := \sum_{j=1}^{n} \lambda_{ij}^{(k)}[\chi, t_0], \quad i = 1, \ldots, n;\ k = 0, 1, 2, \ldots \ . \tag{5}$$

Then, we obtain the following theorem.

Theorem 4 ($\|\cdot\|_\infty$, complex matrix function). *Let $t_0 \in \mathbb{R}_0^+$, let $\chi(t) \in \mathbb{C}^{n\times n}$, $t \geqslant 0$, and $t \mapsto \chi(t)$, $t \geqslant 0$, be $m = 2$ times continuously differentiable, and let the additional condition of Lemma 1 be fulfilled. Further, let $I_{-1} := \{1, \ldots, n\}$ and I_0 be the index set where $\lambda_i^{(0)}[\chi, t_0]$ attains its maximum,*

$$I_0 := \{i_0 \in I_{-1} \mid \lambda_{i_0}^{(0)}[\chi, t_0] = \max_{i \in I_{-1}} \lambda_i^{(0)}[\chi, t_0]\}. \tag{6}$$

Similarly, let

$$I_1 := \{i_1 \in I_0 \mid \lambda_{i_1}^{(1)}[\chi, t_0] = \max_{i \in I_0} \lambda_i^{(1)}[\chi, t_0]\} \tag{7}$$

and

$$I_2 := \{i_2 \in I_1 \mid \lambda_{i_2}^{(2)}[\chi, t_0] = \max_{i \in I_1} \lambda_i^{(2)}[\chi, t_0]\}. \tag{8}$$

Then,

$$\|\chi(t_0)\|_\infty = \max_{i \in I_{-1}} \lambda_i^{(0)}[\chi, t_0], \tag{9}$$

$$D_+^1 \|\chi(t_0)\|_\infty = \max_{i \in I_0} \lambda_i^{(1)}[\chi, t_0], \tag{10}$$

$$D_+^2 \|\chi(t_0)\|_\infty = \max_{i \in I_1} \lambda_i^{(2)}[\chi, t_0]. \tag{11}$$

Remark. If only $m = 1$, then of course still (9) and (10) hold. A similar remark applies in the subsequent theorems and corollaries.

$p = \infty$: *Real $n \times n$ matrix $\chi(t)$*. Define the following sign functionals:

$$s_{ij}^{(0)}[\chi(t_0)] := \operatorname{sgn}[\chi_{ij}(t_0)] \tag{12}$$

and

$$s_{ij}^{(k)}[D^k\chi(t_0)] := \begin{cases} \operatorname{sgn}[\chi_{ij}(t_0)], & \chi_{ij}(t_0) \neq 0, \\ \operatorname{sgn}[(D\chi)_{ij}(t_0)], & \chi_{ij}(t_0) = 0,\ (D\chi)_{ij}(t_0) \neq 0, \\ \operatorname{sgn}[(D^2\chi)_{ij}(t_0)], & \chi_{ij}(t_0) = 0,\ (D\chi)_{ij}(t_0) = 0,\ (D^2\chi)_{ij}(t_0) \neq 0, \\ \vdots & \\ \operatorname{sgn}[(D^k\chi)_{ij}(t_0)], & (D^l\chi)_{ij}(t_0) = 0,\ l = 0,1,\ldots,k-1, \end{cases} \tag{13}$$

$i,j = 1,\ldots,n;\ k = 1,2,\ldots$. This relation can also be written as

$$s_{ij}^{(k)}[D^k\chi(t_0)] = \begin{cases} s_{ij}^{(k-1)}[D^{k-1}\chi(t_0)], & s_{ij}^{(k-1)}[D^{k-1}\chi(t_0)] \neq 0, \\ \operatorname{sgn}[(D^k\chi)_{ij}(t_0)], & s_{ij}^{(k-1)}[D^{k-1}\chi(t_0)] = 0 \end{cases} \tag{14}$$

for $k = 1,2,\ldots$. With these sign functionals, define the further functionals

$$\lambda_i^{(k)}[\chi, t_0] := \sum_{j=1}^{n} s_{ij}^{(k)}[D^k\chi(t_0)]\,(D^k\chi)_{ij}(t_0), \tag{15}$$

$i = 1,\ldots,n;\ k = 0,1,2,\ldots$. Then, the right derivatives for *real matrices* read as follows.

Theorem 5 ($\|\cdot\|_\infty$, real matrix function). *Let $\chi(t) \in \mathbb{R}^{n\times n}$, $t \geqslant 0$, let $t \mapsto \chi(t)$, $t \geqslant 0$, be m times continuously differentiable, and let the additional condition of Lemma* 1 *be fulfilled. Further, let $I_{-1} = \{1,\ldots,n\}$ and I_k be the set of all indices $i_k \in I_{k-1}$, where $\lambda_i^{(k)}[\chi, t_0]$ from* (15) *attains its maximum, i.e.*,

$$I_k := \{i_k \in I_{k-1} \mid \lambda_{i_k}^{(k)}[\chi, t_0] = \max_{i \in I_{k-1}} \lambda_i^{(k)}[\chi, t_0]\}, \tag{16}$$

$k = 0,1,2,\ldots,m.$

Then, the right derivatives of $t \mapsto \|\chi(t)\|_\infty$ at $t = t_0 \geqslant 0$ are given by

$$D_+^k \|\chi(t_0)\|_\infty = \max_{i \in I_{k-1}} \lambda_i^{(k)}[\chi, t_0], \tag{17}$$

$k = 0,1,2,\ldots,m.$

$p = 2$: *Real or complex $n \times n$ matrix $\chi(t)$*. Let $t \mapsto \chi(t)$, $t \geqslant 0$, be analytic for ease of treatment. We mention, however, that the formulae to be derived will remain valid if this function is only $m = 2$ times continuously differentiable and if the additional condition of Lemma 1 is fulfilled. Starting point in the case $p = 2$ is the series expansion

$$P(t) := \chi^*(t)\,\chi(t) = \sum_{j=0}^{\infty} \beta_j \frac{(t-t_0)^j}{j!}, \quad t \geqslant t_0, \tag{18}$$

with

$$\beta_j = \sum_{k=0}^{j} \binom{j}{k} D^k\chi^*(t_0)\, D^{j-k}\chi(t_0), \quad j = 0,1,2,\ldots \ . \tag{19}$$

Now, define

$$T^{(0)} = \begin{cases} \chi^*(t_0)\,\chi(t_0), & \chi(t_0) \neq 0, \\ D\chi^*(t_0)\,D\chi(t_0), & \chi(t_0) = 0,\ D\chi(t_0) \neq 0, \\ \frac{1}{4}\,D^2\chi^*(t_0)\,D^2\chi(t_0), & \chi(t_0) = 0,\ D\chi(t_0) = 0, \end{cases} \tag{20}$$

$$T^{(1)} = \begin{cases} \chi^*(t_0)\,D\chi(t_0) + D\chi^*(t_0)\,\chi(t_0), & \chi(t_0) \neq 0, \\ \frac{1}{2}\,[D\chi^*(t_0)\,D^2\chi(t_0) + D^2\chi^*(t_0)\,D\chi(t_0)], & \chi(t_0) = 0, \end{cases} \tag{21}$$

$$T^{(2)} = \tfrac{1}{2}\,[\chi^*(t_0)\,D^2\chi(t_0) + 2\,D\chi^*(t_0)\,D\chi(t_0) + D^2\chi^*(t_0)\,\chi(t_0)]. \tag{22}$$

Then,

$$P(t) = \begin{cases} T^{(0)} + T^{(1)}\,(t-t_0) + T^{(2)}\,(t-t_0)^2 + \cdots, & \chi(t_0) \neq 0, \\ (t-t_0)^2\,[T^{(0)} + T^{(1)}\,(t-t_0) + \cdots], & \chi(t_0) = 0,\ D\chi(t_0) \neq 0, \\ (t-t_0)^4\,[T^{(0)} + \cdots], & \chi(t_0) = 0,\ D\chi(t_0) = 0. \end{cases} \tag{23}$$

Let $\lambda_{\max}(P(t))$ be the largest eigenvalue of $P(t)$. Then, due to [10, Theorem 5.11, Chapter II, pp. 115–116] and [11, Lemma 2.1],

$$\|\chi(t)\|_2 = \begin{cases} [v_0 + v_1\,(t-t_0) + v_2\,(t-t_0)^2 + \cdots]^{1/2}, & \chi(t_0) \neq 0, \\ (t-t_0)\,[v_0 + v_1\,(t-t_0) + \cdots]^{1/2}, & \chi(t_0) = 0,\ D\chi(t_0) \neq 0, \\ (t-t_0)^2\,[v_0 + \cdots]^{1/2}, & \chi(t_0) = 0,\ D\chi(t_0) = 0, \end{cases} \tag{24}$$

$t \geqslant t_0$, where $v_0,\ v_1,\ v_2$ are defined in [12, (33), (35), (38)], as the case may be, since $\|\chi(t)\|_2 = [\lambda_{\max}(P(t))]^{1/2}$. Hereby, we obtain

Theorem 6 ($\|\cdot\|_2$, real or complex matrix function). *Let $\chi(t) \in \mathbb{C}^{n\times n}$, $t \geqslant 0$, let $t \mapsto \chi(t)$, $t \geqslant 0$, be $m = 2$ times continuously differentiable, and let the additional condition of Lemma* 1 *be fulfilled. Further, let $T^{(0)}$, $T^{(1)}$, and $T^{(2)}$ be defined by* (20), (21), *and* (22), *as well as v_0, v_1, v_2 by* [12, (33), (35), (38)], *as the case may be.*

Then,

$$\|\chi(t_0)\|_2 = \begin{cases} v_0^{1/2}, & \chi(t_0) \neq 0, \\ 0, & \chi(t_0) = 0, \end{cases} \tag{25}$$

$$D_+^1\|\chi(t_0)\|_2 = \begin{cases} \dfrac{1}{2}\,\dfrac{v_1}{v_0^{1/2}}, & \chi(t_0) \neq 0, \\ v_0^{1/2}, & \chi(t_0) = 0,\ D\chi(t_0) \neq 0, \\ 0, & \chi(t_0) = 0,\ D\chi(t_0) = 0, \end{cases} \tag{26}$$

$$D_+^2\|\chi(t_0)\|_2 = \begin{cases} \dfrac{1}{2}\,\dfrac{2\,v_0\,v_2 - \frac{1}{2}\,v_1^2}{v_0^{3/2}}, & \chi(t_0) \neq 0, \\ \dfrac{v_1}{v_0^{1/2}}, & \chi(t_0) = 0,\, D\chi(t_0) \neq 0, \\ 2\,v_0^{1/2}, & \chi(t_0) = 0,\, D\chi(t_0) = 0,\, D^2\chi(t_0) \neq 0, \\ 0, & \chi(t_0) = 0,\, D\chi(t_0) = 0,\, D^2\chi(t_0) = 0. \end{cases} \tag{27}$$

Remark. The reader should notice that the quantity v_0 in Formula (26) for $\chi(t_0) \neq 0$ is different from that for $\chi(t_0) = 0$, $D\chi(t_0) = 0$. A similar remark holds for Formula (27).

3.2. Matrix functions $t \mapsto \chi(t)$ in the norms $|\cdot|_p$, $1 < p < \infty$

For the p-norms $|\cdot|_p$, $1 < p < \infty$, new results similar to those in [13] for complex-valued vector functions are obtained.

$1 < p < \infty$: *Complex* $n \times n$ *matrix* $\chi(t)$. Let $t \mapsto \chi(t) \in \mathbb{C}^{n\times n}$ be $m = 2$ times continuously differentiable, and

$$\chi(t) = \chi(t_0) + (t - t_0)\,D\chi(t_0) + \frac{(t-t_0)^2}{2!}\,D^2\chi(t_0) + \rho(t), \quad t \geqslant t_0,$$

where

$$\rho(t) = \mathrm{O}((t-t_0)^2).$$

With these matrices, define the following functionals for $i, j \in \{1, \ldots, n\}$:

$$\chi_{ij}^{(0)} := |\chi_{ij}(t_0)|, \tag{28}$$

$$\chi_{ij}^{(1)} := \begin{cases} \dfrac{\mathrm{Re}\,\chi_{ij}(t_0)\,\mathrm{Re}\,(D\chi)_{ij}(t_0) + \mathrm{Im}\,\chi_{ij}(t_0)\,\mathrm{Im}\,(D\chi)_{ij}(t_0)}{|\chi_{ij}(t_0)|}, & \chi_{ij}(t_0) \neq 0, \\ |(D\chi)_{ij}(t_0)|, & \chi_{ij}(t_0) = 0, \end{cases} \tag{29}$$

$$\chi_{ij}^{(2)} := \begin{cases} \frac{|(D\chi)_{ij}(t_0)|^2 + \mathrm{Re}\,\chi_{ij}(t_0)\,\mathrm{Re}\,(D^2\chi)_{ij}(t_0) + \mathrm{Im}\,\chi_{ij}(t_0)\,\mathrm{Im}\,(D^2\chi)_{ij}(t_0)}{|\chi_{ij}(t_0)|} & \\ \quad - \frac{[\mathrm{Re}\,\chi_{ij}(t_0)\,\mathrm{Re}\,(D\chi)_{ij}(t_0) + \mathrm{Im}\,\chi_{ij}(t_0)\,\mathrm{Im}\,(D\chi)_{ij}(t_0)]^2}{|\chi_{ij}(t_0)|^3}, & \chi_{ij}(t_0) \neq 0, \\ \frac{\mathrm{Re}\,(D\chi)_{ij}(t_0)\,\mathrm{Re}\,(D^2\chi)_{ij}(t_0) + \mathrm{Im}\,(D\chi)_{ij}(t_0)\,\mathrm{Im}\,(D^2\chi)_{ij}(t_0)}{|(D\chi)_{ij}(t_0)|}, & \chi_{ij}(t_0) = 0,\ (D\chi)_{ij}(t_0) \neq 0, \\ |(D^2\chi)_{ij}(t_0)|, & \chi_{ij}(t_0) = 0,\ (D\chi)_{ij}(t_0) = 0. \end{cases} \tag{30}$$

Hereby, define the further functionals:

$$\chi^{(0,p)} := \left(\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^p\right)^{1/p}, \tag{31}$$

$$\chi^{(1,p)} := \begin{cases} \dfrac{\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p-1}\, \chi_{ij}^{(1)}}{(\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p})^{1-1/p}}, & \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p}\right)^{1/p} \neq 0, \\ \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(1)})^{p}\right)^{1/p}, & \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p}\right)^{1/p} = 0, \end{cases} \tag{32}$$

$$\chi^{(2,p)} := \begin{cases} \dfrac{\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p-1}\, \chi_{ij}^{(2)} + (p-1)\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p-2} (\chi_{ij}^{(1)})^{2}}{(\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p})^{1-1/p}}, & \\ \quad + \dfrac{(1-p)[\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p-1}\, \chi_{ij}^{(1)}]^{2}}{(\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p})^{2-1/p}}, & \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p}\right)^{1/p} \neq 0, \\ \dfrac{\sum_{i,j=1}^{n} (\chi_{ij}^{(1)})^{p-1}\, \chi_{ij}^{(2)}}{(\sum_{i,j=1}^{n} (\chi_{ij}^{(1)})^{p})^{1-1/p}}, & \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p}\right)^{1/p} = 0, \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(1)})^{p}\right)^{1/p} \neq 0, \\ \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(2)})^{p}\right)^{1/p}, & \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(0)})^{p}\right)^{1/p} = 0, \left(\displaystyle\sum_{i,j=1}^{n} (\chi_{ij}^{(1)})^{p}\right)^{1/p} = 0. \end{cases} \tag{33}$$

Then, we obtain

Theorem 7 ($|\cdot|_p$, $1 < p < \infty$, complex matrix function). *Let $\chi : \mathbb{R}_0^+ \to \mathbb{C}^{n\times n}$ be a matrix function that is $m = 2$ times continuously differentiable, and let $t_0 \in \mathbb{R}_0^+$.*

Then,

$$|\chi(t_0)|_p = \chi^{(0,p)}, \tag{34}$$

$$D_+^1 |\chi(t_0)|_p = \chi^{(1,p)}, \tag{35}$$

$$D_+^2 |\chi(t_0)|_p = \chi^{(2,p)}, \tag{36}$$

where $\chi^{(0,p)}$-$\chi^{(2,p)}$ are defined by (31)–(33).

Special case: $|\cdot|_2$. In the special case $p = 2$, it is useful to introduce a scalar product in the set of complex $n \times n$ matrices. So, let $B = (b_{ij})$, $C = (c_{ij}) \in \mathbb{C}^{n\times n}$ and define

$$(B, C) := \sum_{i,j=1}^{n} b_{ij}\, \bar{c}_{ij}. \tag{37}$$

Then, $(\cdot,\cdot)$ defines a scalar product in $\mathbb{C}^{n\times n}$ and

$$|B|_2 = (B, B)^{1/2}. \tag{38}$$

Hereby, the right derivatives in the norm $|\cdot|_2$ can also be written as follows (cf. [13]):

$$D_+^1|\chi(t_0)|_2 = \begin{cases} \dfrac{\mathrm{Re}\,(\chi(t_0), D\chi(t_0))}{|\chi(t_0)|_2}, & \chi(t_0) \neq 0, \\ |D\chi(t_0)|_2, & \chi(t_0) = 0, \end{cases} \tag{39}$$

$$D_+^2|\chi(t_0)|_2 = \begin{cases} \dfrac{|D\chi(t_0)|_2^2 + \mathrm{Re}\,(\chi(t_0), D^2\chi(t_0))}{|\chi(t_0)|_2} - \dfrac{[\mathrm{Re}\,(\chi(t_0), D\chi(t_0))]^2}{|\chi(t_0)|_2^3}, & \chi(t_0) \neq 0, \\ \dfrac{\mathrm{Re}\,(D\chi(t_0), D^2\chi(t_0))}{|D\chi(t_0)|_2}, & \chi(t_0) = 0,\ D\chi(t_0) \neq 0, \\ |D^2\chi(t_0)|_2, & \chi(t_0) = 0,\ \chi(t_0) = 0. \end{cases} \tag{40}$$

Logarithmic derivatives in the norms $|\cdot|_p,\ 1 < p < \infty$. Let $A \in \mathbb{C}^{n\times n}$. Then, the two first logarithmic derivatives $\mu_p^{(k)}[A],\ k = 1, 2$ are defined by

$$\mu_p^{(k)}[A] = D_+^{(k)}|\Phi(0)|_p, \quad k = 1, 2. \tag{41}$$

So, one has to set $t_0 = 0$ and

$$\chi(t_0) = E,$$
$$D\chi(t_0) = A,$$
$$D^2\chi(t_0) = A^2.$$

Since $E = (\delta_{ij})$ is a diagonal matrix, the double sum over the index pairs ij reduces to a simple sum over the index pairs ii, in the sequel. Thus, due to (28)–(30) and (32), (33) as well as $|A_{ii}|^2 = (\mathrm{Re}\,A_{ii})^2 + (\mathrm{Im}\,A_{ii})^2,\ i = 1, \ldots, n$, the following formulae are obtained:

$$\mu_p^{(1)}[A] = \frac{1}{n^{1-1/p}} \left\{ \sum_{i=1}^n \mathrm{Re}\,A_{ii} \right\}, \tag{42}$$

resp., with the eigenvalues $\lambda_i(A),\ i = 1, \ldots, n$ of A,

$$\mu_p^{(1)}[A] = \frac{1}{n^{1-1/p}} \left\{ \sum_{i=1}^n \mathrm{Re}\,\lambda_i(A) \right\} \tag{43}$$

and

$$\mu_p^{(2)}[A] = \begin{cases} \dfrac{1}{n^{1-1/p}} \left\{ \displaystyle\sum_{i=1}^n \mathrm{Re}\,(A^2)_{ii} + \sum_{i=1}^n (\mathrm{Im}\,A_{ii})^2 + (p-1)\sum_{i=1}^n (\mathrm{Re}\,A_{ii})^2 \right\} \\ \quad + \dfrac{1-p}{n^{2-1/p}} \left(\displaystyle\sum_{i=1}^n \mathrm{Re}\,A_{ii} \right)^2, & p \neq 2, \\ \dfrac{1}{n^{1/2}} \left\{ \displaystyle\sum_{i=1}^n \mathrm{Re}\,(A^2)_{ii} + |A|_2^2 \right\} - \dfrac{1}{n^{3/2}} \left(\displaystyle\sum_{i=1}^n \mathrm{Re}\,A_{ii} \right)^2, & p = 2. \end{cases} \tag{44}$$

Remark. In Formula (44), one can substitute

$$\sum_{i=1}^{n} \operatorname{Re} A_{ii} = \sum_{i=1}^{n} \operatorname{Re} \lambda_i(A), \quad \sum_{i=1}^{n} \operatorname{Re} (A^2)_{ii} = \sum_{i=1}^{n} \operatorname{Re} [\lambda_i(A)]^2,$$

$$|A|_2^2 = \operatorname{trace}(A^*A) = \sum_{i=1}^{n} \lambda_i(A^*A).$$

Further, $\lim_{\substack{p \neq 2 \\ p \to 2}} \mu_p^{(2)}[A] \neq \mu_2^{(2)}[A]$. The reason for this is that $p=2$ plays a special role in Formula (33). The details are left to the reader. We mention that Formula (44) for $p=2$ can be checked by Formula (40).

4. Upper bounds on some matrix functions

We first consider general matrix power functions, then turn to the discrete evolution, and finally to the remainder function.

4.1. General matrix power functions

Let $B \in \mathbb{C}^{n\times n}$ be a matrix. For later use, we want to derive an upper bound on the matrix power function $x \mapsto B^x$, $x \geqslant 0$. Starting point of our investigation is the subsequent reformulation of a well-known result for the matrix exponential $t \mapsto e^{At}$, $t \geqslant 0$, where $A \in \mathbb{C}^{n\times n}$ is a given matrix. By $\sigma(A)$, we denote the spectrum of A, that is, the set of all eigenvalues of A, by $v(A)$ the spectral abscissa of A, that is, $v(A) = \max_{\lambda\in\sigma(A)} \operatorname{Re}\lambda$, and by $\rho(A)$ the spectral radius of A, that is, $\rho(A) = \max_{\lambda\in\sigma(A)} |\lambda|$. The index $i(\lambda)$ of an eigenvalue $\lambda \in \sigma(A)$ is defined as the maximal dimension of the corresponding Jordan blocks of matrix A (cf. [4, p. 76]).

Lemma 8 (Reformulation of a well-known result). *Let $\|\cdot\|$ be any matrix norm, let $A \in \mathbb{C}^{n\times n}$, and let $\rho(e^A)$ be the spectral radius of e^A.*

Then, for every $\varepsilon > 0$ there exists a constant $M_\varepsilon > 0$ such that

$$\|e^{At}\| \leqslant M_\varepsilon (\rho(e^A) + \varepsilon)^t, \quad t \geqslant 0. \tag{45}$$

If, additionally, for every eigenvalue $\mu \in \sigma(e^A)$ with $|\mu| = \rho(e^A)$ the index $i(\mu)$ of μ satisfies $i(\mu) = 1$, then the above bound is also valid for $\varepsilon = 0$.

Proof. Let the conditions of the Lemma be fulfilled. Then, it is well-known that

$$\|e^{At}\| \leqslant M_\varepsilon e^{(v(A)+\varepsilon)t}, \quad t \geqslant 0 \tag{46}$$

(see [4, p. 78]). If, additionally, for every eigenvalue $\lambda \in \sigma(A)$ with $\operatorname{Re}(\lambda) = v(A)$ the index $i(\lambda)$ of λ satisfies $i(\lambda) = 1$, then the bound (46) is also valid for $\varepsilon = 0$.

Now, let $\lambda_j = \lambda_j(A)$, $j = 1,\dots,n$ be the eigenvalues of A. Then, $\mu_j = \mathrm{e}^{\lambda_j}$, $j = 1,\dots,n$ are the eigenvalues of e^A and

$$\begin{aligned} e^{\nu(A)} &= e^{\max_{j=1,\dots,n} \operatorname{Re}\lambda_j} = \max_{j=1,\dots,n} e^{\operatorname{Re}\lambda_j} \\ &= \max_{j=1,\dots,n} |e^{\lambda_j}| = \max_{j=1,\dots,n} |\lambda_j(e^A)| = \max_{j=1,\dots,n} |\mu_j| \\ &= \rho(e^A). \end{aligned}$$

Thus, according to (46),

$$\|e^{A\,t}\| \leqslant M_\varepsilon\,(e^\varepsilon \cdot \rho(e^A))^t, \quad t \geqslant 0.$$

Let $\varepsilon > 0$ and let $\tilde\varepsilon = \tilde\varepsilon(\varepsilon) > 0$ be such that

$$e^{\tilde\varepsilon} = 1 + \frac{\varepsilon}{\rho(e^A)}.$$

Then, there exists a constant $\tilde M_{\tilde\varepsilon}(\varepsilon) > 0$ such that

$$\begin{aligned} \|e^{A\,t}\| &\leqslant \tilde M_{\tilde\varepsilon}(\varepsilon)\,(e^{\tilde\varepsilon} \cdot \rho(e^A))^t \\ &= \tilde M_{\tilde\varepsilon}(\varepsilon)\left(\left(1 + \frac{\varepsilon}{\rho(e^A)}\right)\,\rho(e^A)\right)^t \\ &= \tilde M_{\tilde\varepsilon}(\varepsilon)\,(\rho(e^A) + \varepsilon)^t, \quad t \geqslant 0. \end{aligned}$$

Set $M_\varepsilon = \tilde M_{\tilde\varepsilon}(\varepsilon)$. Then, the assertion follows.

The additional condition for $\lambda \in \sigma(A)$ is fulfilled if and only if the additional condition for $\mu = \mathrm{e}^\lambda \in \sigma(e^A)$ is satisfied, that is, one has $i(\lambda) = 1$ if and only if $i(\mu) = 1$. This follows from the associated Jordan forms. So, if the additional condition is fulfilled, $\varepsilon = 0$ can be chosen. □

For $x \in \mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, \dots\}$, B^x is well-defined. For arbitrary $x \geqslant 0$, let $0 \notin \sigma(B)$. Then, we define (cf. [1, pp. 38–40])

$$B^x := e^{x \ln B}, \quad x \geqslant 0. \tag{47}$$

From Lemma 8, it is clear what we can expect to prove for the matrix power function $x \mapsto B^x$, $x \geqslant 0$. We obtain

Lemma 9. *Let $\|\cdot\|$ be any matrix norm, let $B \in \mathbb{C}^{n\times n}$, and let $0 \notin \sigma(B)$. Then, for every $\varepsilon > 0$ there exists a constant $M_\varepsilon > 0$ such that*

$$\|B^x\| \leqslant M_\varepsilon\,(\rho(B) + \varepsilon)^x, \quad x \geqslant 0. \tag{48}$$

If, additionally, for every eigenvalue $\mu \in \sigma(B)$ with $|\mu| = \rho(B)$ the index $i(\mu)$ of μ satisfies $i(\mu) = 1$, then the above bound is also valid for $\varepsilon = 0$.

Proof. Set $A = \ln B$. Then, by Lemma 8,

$$\|B^x\| = \|e^{A\,x}\| \leqslant M_\varepsilon\,(\rho(e^A) + \varepsilon)^x = M_\varepsilon\,(\rho(B) + \varepsilon)^x, \quad x \geqslant 0.$$

Further, the eigenvalues λ resp. μ of A resp. B are related by $\lambda = \ln \mu$, and $i(\lambda) = 1$ if and only if $i(\mu) = 1$. So, $\varepsilon = 0$ can be set if the additional condition is fulfilled. $\square$

Remark. In case of $B = e^A$ with a given matrix A, the condition $0 \notin \sigma(B)$ is automatically satisfied. Further, if $x \in \mathbb{N}_0$, the estimate (48) is well-known (cf. [4, p. 90] for an equivalent representation). $\square$

4.2. Discrete evolution

The solution to the initial-value problem $\dot{x}(t) = A\,x(t)$, $x(0) = x_0$, is given by $x(t) = \Phi(t)\,x_0$ with $\Phi(t) = e^{A\,t}$. In this context, $\Phi(t)$ is called fundamental matrix or evolution.

In the numerical approximation of the solution $x(t)$ at the grid points $t_r = r\,\Delta\,t$, $r = 0, 1, 2, \ldots$ with given $\Delta t > 0$, the discrete equivalent $\Psi_{\Delta t}$ of $\Phi(\Delta t)$ is encountered.

Every finite difference method with order k and k stages can be written in the form

$$x_{r+1} = \Psi_{\Delta t}\,x_r, \quad r = 0, 1, 2, \ldots \tag{49}$$

so that x_r is an approximation of $x(t_r) = x(r\,\Delta t)$, $r = 1, 2, \ldots$. The solution of (49) is given by

$$x_r = (\Psi_{\Delta t})^r\,x_0, \quad r = 0, 1, 2, \ldots \ . \tag{50}$$

Consequently, the power function $(\Psi_{\Delta t})^r$ is the discrete equivalent to the (continuous) evolution $\Phi(t) = \Phi(r\,\Delta t) = [\Phi(\Delta t)]^r$, $r = 0, 1, 2, \ldots$. We remark that r in $(\Psi_{\Delta t})^r$ and t in $\Phi(t)$ are related by

$$r = \frac{t}{\Delta t}. \tag{51}$$

Now, we define the powers $(\Psi_{\Delta t})^r$ for all $r \geqslant 0$ resp. $(\Psi_{\Delta t})^{t/\Delta t}$ for all $t \geqslant 0$. For this, let $0 \notin \sigma(\Psi_{\Delta t})$. Then, according to (45), the discrete evolution $\Psi(t)$, $t \geqslant 0$, is well defined by

$$\Psi(t) := (\Psi_{\Delta t})^{t/\Delta t}, \quad t \geqslant 0. \tag{52}$$

One can use this extension to $\mathbb{R}_0^+$, for example, to compute the approximate values $x_{r+\delta}$ of $x((r+\delta)\,\Delta t)$ with $0 < \delta < 1$ for fixed Δt.

Remark. The referee has pointed out how $x_{r+\delta}$ can be approximated by using only integer powers r of $\Psi_{\Delta t}$ (see Section 5.4). $\square$

Special cases for $\Psi_{\Delta t}$ are the *explicit Euler method*, when

$$\Psi_{\Delta t} = E + A\,\Delta t \tag{53}$$

or the *Runge–Kutta method*, when

$$\Psi_{\Delta t} = E + A\,\Delta t + A^2\,\frac{(\Delta t)^2}{2!} + A^3\,\frac{(\Delta t)^3}{3!} + A^4\,\frac{(\Delta t)^4}{4!}, \tag{54}$$

or, more generally, each *partial sum of* $e^{A\,t}$, when

$$\Psi_{\Delta t} = \sum_{i=0}^{k} A^i\,\frac{(\Delta t)^i}{i!}, \tag{55}$$

where $k \in \mathbb{N},\ k \geqslant 2$. Then, the condition $0 \notin \sigma(\Psi_{\Delta t})$ is equivalent to the *condition*

$$\text{(C)} \quad \sum_{i=0}^{k} \lambda^i \frac{(\Delta t)^i}{i!} \neq 0, \quad \lambda \in \sigma(A),$$

which is fulfilled if $\Delta t > 0$ is sufficiently small.

From Lemma 9, we obtain

Corollary 10. *Let $\|\cdot\|$ be any matrix norm. Let $A \in \mathbb{C}^{n\times n}$, let $k \in \mathbb{N},\ k \geqslant 2$, and let $0 \notin \sigma(\Psi_{\Delta t})$. Further, let $\Psi_{\Delta t}$ be defined by (55) and Ψ by (52).*

Then, for every $\varepsilon > 0$ there exists a constant $M_\varepsilon > 0$ such that

$$\|\Psi(t)\| = \|(\Psi_{\Delta t})^{t/\Delta t}\| \leqslant M_\varepsilon\,[\rho(\Psi_{\Delta t}) + \varepsilon]^{t/\Delta t}, \quad t \geqslant 0. \tag{56}$$

If, additionally, for every eigenvalue $\mu \in \sigma(\Psi_{\Delta t})$ with $|\mu| = \rho(\Psi_{\Delta t})$ the index $i(\mu)$ is equal to $i(\mu) = 1$, then $\varepsilon = 0$ can be chosen.

4.3. Difference between evolution and discrete evolution

Define

$$R(t) := \Phi(t) - \Psi(t), \quad t \geqslant 0. \tag{57}$$

For this difference (or remainder), we want to obtain an upper bound. This is based on the following lemma.

Lemma 11. *Let the eigenvalues of A be simple and X be the modal matrix of A, i.e., $X = [x_1, x_2, \ldots, x_n]$, where the x_i are the linearly independent eigenvectors associated with the eigenvalues $\lambda_i,\ i = 1, \ldots, n$ of A. If $f(z)$ and $g(z)$ are analytic on an open set containing the spectrum $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$ of A, then*

$$\|f(A) - g(A)\|_2 \leqslant \kappa_2(X) \max_{i=1,\ldots,n} |f(\lambda_i) - g(\lambda_i)|, \tag{58}$$

where $\kappa_2(X) = \|X\|_2\,\|X^{-1}\|_2$ is the condition number in the spectral norm.

Proof. This lemma follows from [5, p. 547, Theorem 11.2.1] with $p = n$ and $m_i = 1,\ i = 1, \ldots, n$. □

Now, let

$$f(z) := e^{z\,t}$$

and

$$g(z) := s(z) := \left(\sum_{i=0}^{k} z^i \frac{(\Delta t)^i}{i!} \right)^{t/\Delta t}.$$

Then,

$$f(A) = e^{A\,t} = \Phi(t)$$

and

$$g(A) = \left(\sum_{i=0}^{k} A^i \, \frac{(\Delta t)^i}{i!}\right)^{t/\Delta t} = (\Psi_{\Delta t})^{t/\Delta t} = \Psi(t).$$

Define

$$r_j(t) := \mathrm{e}^{\lambda_j\, t} - \left(\sum_{i=0}^{k} \lambda_j^i \, \frac{(\Delta t)^i}{i!}\right)^{t/\Delta t}, \quad j = 1, \ldots, n \tag{59}$$

as well as

$$r(t) = [r_1(t), \ldots, r_n(t)]^{\mathrm{T}}. \tag{60}$$

Then, we obtain

Corollary 12. *Let the eigenvalues of A be simple, let X be the associated modal matrix, and let the condition (C) be fulfilled. Then,*

$$\|e^{A\,t} - (\Psi_{\Delta t})^{t/\Delta t}\|_2 \leqslant \kappa_2(X) \max_{i=1,\ldots,n} \left| \mathrm{e}^{\lambda_j\, t} - \left(\sum_{i=0}^{k} \lambda_j^i \, \frac{(\Delta t)^i}{i!}\right)^{t/\Delta t} \right|, \quad t \geqslant 0, \tag{61}$$

or, with $R(t)$ in (57) and $r(t)$ in (60),

$$\|R(t)\|_2 \leqslant \kappa_2(X)\, \|r(t)\|_\infty, \quad t \geqslant 0. \tag{62}$$

Remark. In the application part, we need the first derivative of $r(t)$, which is given by

$$D^1 r_j(t) = \lambda_j \, \mathrm{e}^{\lambda_j\, t} - \frac{\ln\{(\sum_{i=0}^{k} \lambda_j^i \, \frac{(\Delta t)^i}{i!})^{t/\Delta t}\}}{\Delta t} \left(\sum_{i=0}^{k} \lambda_j^i \, \frac{(\Delta t)^i}{i!}\right)^{t/\Delta t} \quad j = 1, \ldots, n.$$

5. Applications

In this section, we apply the obtained results to a vibration problem and get the best upper bounds in certain classes of upper bounds for $\chi \in \mathscr{F} = \{\Psi, R\}$. This is achieved by combining the differential calculus of norms, developed in this paper, and upper bounds, obtained by classical methods. The results are illustrated by graphics. Beyond this, also some numerical values are given in order that the reader may check and compare the computations.

5.1. Multi-mass vibration problem

We take up the multi-mass vibration model of [12] shown in Fig. 1.

The associated initial-value problem is given by

$$M\,\ddot{y} + B\,\dot{y} + K\,y = 0, \quad y(0) = y_0, \; \dot{y}(0) = \dot{y}_0,$$

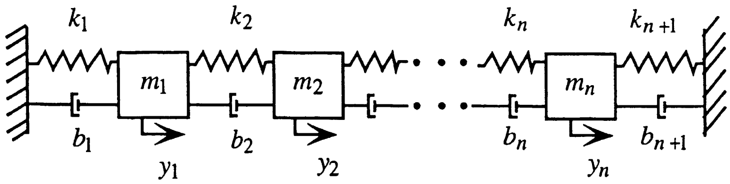

Fig. 1. Multi-mass vibration model.

with the matrices M, B, K and the displacement vector y as in [12]. In state–space description, this problem takes the form

$$\dot{x}(t) = A\,x(t), \quad x(0) = x_0,$$

where $x = [y^{\mathrm{T}}, z^{\mathrm{T}}]^{\mathrm{T}}$, $z = \dot{y}$, and where the system matrix A is given by

$$A = \left[\begin{array}{c|c} 0 & E \\ \hline -M^{-1}K & -M^{-1}B \end{array}\right].$$

The values for m_j, $j = 1, \ldots, n$ and for b_j, k_j, $j = 1, \ldots, n+1$ are also specified as in [12]. Moreover, we choose the stepsize $\mathbf{\Delta t = 0.1}$ as well as the summation index $\mathbf{k = 4}$ in $\Psi_{\Delta t} = \sum_{i=0}^{k} A^i\,\Delta t/i!$, i.e., the Runge–Kutta method. Further, $\mathbf{n = 5}$ is set so that the state–space vector $x(t)$ has the dimension $m = 2\,n = 10$. Again, we choose $\varepsilon = eps = 2^{-52} \doteq 2.2204 \times 10^{-16}$ (the machine precision of MATLAB).

5.2. Optimal upper bounds on the discrete evolution based on Corollary 10

Let $\|\cdot\|$ be any matrix norm for which $t \mapsto \|\Psi(t)\|$ is sufficiently regular. To obtain the minimal M_ε such that $\|\Psi(t)\| \leqslant M_\varepsilon\,(\rho(\Psi_{\Delta t}) + \varepsilon)^{t/\Delta t}$, $t \geqslant 0$, we seek a place t_c where the function

$$t \mapsto \psi_{M_\varepsilon}(t) := M_\varepsilon\,(\rho(\Psi_{\Delta t}) + \varepsilon)^{t/\Delta t}, \quad t \geqslant 0,$$

meets the function $t \mapsto \|\Psi(t)\|$. Thus,

$$\|\Psi(t_c)\| \stackrel{!}{=} \psi_{M_\varepsilon}(t_c)$$

and

$$D_+^1 \|\Psi(t_c)\| \stackrel{!}{=} \psi'_{M_\varepsilon}(t_c) = \frac{\ln(\rho(\Psi_{\Delta t}) + \varepsilon)}{\Delta t}\,\psi_{M_\varepsilon}(t_c).$$

This is a system of two nonlinear equations in the two unknowns t_c and M_ε. By eliminating $\psi_{M_\varepsilon}(t_c)$, this system is reduced to

$$D_+^1 \|\Psi(t_c)\| = \frac{\ln(\rho(\Psi_{\Delta t}) + \varepsilon)}{\Delta t}\,\|\Psi(t_c)\|,$$

which is a single nonlinear equation in the single unknown t_c.

The results for $\|\cdot\| = \|\cdot\|_p$ with $p \in \{\infty, 2\}$ are similar to those for $\|\Phi(t)\|_p$ (cf. [12, Figs. 2 and 5]). Therefore, the corresponding figures are not shown.

5.3. Optimal upper bounds on the difference between the evolution and the discrete evolution based on Corollary 12

In view of Corollary 12 and the equivalence of norms, there are constants $f_{2,\infty}=1$, $f_{2,2}$, $f_{\infty,\infty}$, and $\tilde{f}_{2,2}$ such that

$$\|R(t)\|_2 \leqslant f_{2,\infty}\,\kappa_2(X)\,\|r(t)\|_\infty, \quad t \geqslant 0, \tag{63}$$

$$\|R(t)\|_2 \leqslant f_{2,2}\,\kappa_2(X)\,\|r(t)\|_2, \quad t \geqslant 0, \tag{64}$$

$$\|R(t)\|_\infty \leqslant f_{\infty,\infty}\,\kappa_2(X)\,\|r(t)\|_\infty, \quad t \geqslant 0, \tag{65}$$

$$|R(t)|_2 \leqslant \tilde{f}_{2,2}\,\kappa_2(X)\,\|r(t)\|_2, \quad t \geqslant 0. \tag{66}$$

In this subsection, we shall compute the associated optimal values $f^*_{2,\infty}$, $f^*_{2,2}$, $f^*_{\infty,\infty}$, and $\tilde{f}^*_{2,2}$ by applying the new results of the differential calculus for norms of the matrix function $\chi=R$.

(i) *Case* $\|R(t)\|_2$ *and* $\|r(t)\|_\infty$: For example, to obtain the minimal constant $f_{2,\infty}=f^*_{2,\infty}$ such that (63) holds, we seek a place $t^*_{2,\infty}$, where the function

$$t \mapsto \|R(t)\|_2, \quad t \geqslant 0,$$

meets the function

$$t \mapsto f^*_{2,\infty}\,\kappa_2(X)\,\|r(t)\|_\infty, \quad t \geqslant 0.$$

Thus,

$$\|R(t^*_{2,\infty})\|_2 \stackrel{!}{=} f^*_{2,\infty}\,\kappa_2(X)\,\|r(t^*_{2,\infty})\|_\infty \tag{67}$$

and

$$D^1_+\,\|R(t^*_{2,\infty})\|_2 \stackrel{!}{=} f^*_{2,\infty}\,\kappa_2(X)\,D^1_+\,\|r(t^*_{2,\infty})\|_\infty. \tag{68}$$

This is a system of two nonlinear equations in the two unknowns $t^*_{2,\infty}$ and $f^*_{2,\infty}$. By eliminating $f^*_{2,\infty}$, this system is reduced to

$$\frac{D^1_+\,\|R(t^*_{2,\infty})\|_2}{\|R(t^*_{2,\infty})\|_2} = \frac{D^1_+\,\|r(t^*_{2,\infty})\|_\infty}{\|r(t^*_{2,\infty})\|_\infty}$$

or

$$D^1_+\,\|R(t^*_{2,\infty})\|_2\,\|r(t^*_{2,\infty})\|_\infty - \|R(t^*_{2,\infty})\|_2\,D^1_+\,\|r(t^*_{2,\infty})\|_\infty = 0, \tag{69}$$

which is a simple nonlinear equation in the single unknown $t^*_{2,\infty}$. When $t^*_{2,\infty}$ has been computed from (69), $f^*_{2,\infty}$ is obtained from

$$f^*_{2,\infty} = \frac{\|R(t^*_{2,\infty})\|_2}{\kappa_2(X)\,\|r(t^*_{2,\infty})\|_\infty}. \tag{70}$$

For the given data, we obtain

$$\lambda_1 \doteq -0.69976063878054 + 1.7959814781598\,i,$$

$$\lambda_2 \doteq -0.69976063878054 - 1.7959814781598\,i,$$

$$\lambda_3 \doteq -0.56266837404074 + 1.6163587016439\, i,$$

$$\lambda_4 \doteq -0.56266837404074 - 1.6163587016439\, i,$$

$$\lambda_5 \doteq -0.375 + 1.3635890143295\, i,$$

$$\lambda_6 \doteq -0.375 - 1.3635890143295\, i,$$

$$\lambda_7 \doteq -0.18733162595926 + 0.99452168646559\, i,$$

$$\lambda_8 \doteq -0.18733162595926 - 0.99452168646559\, i,$$

$$\lambda_9 \doteq -0.050239361219464 + 0.51637145071101\, i,$$

$$\lambda_{10} \doteq -0.050239361219464 - 0.51637145071101\, i$$

and

$$\kappa_2(X) \doteq 2.48602721717244.$$

Further, (69) and (70) deliver

$$t^*_{2,\infty} \doteq 4.99800544343835,$$

$$f^*_{2,\infty} \doteq 0.89912238425724,$$

$$M^*_{2,\infty} := f^*_{2,\infty}\, \kappa_2(X) \doteq 2.23524271881451.$$

These values and similar values, that follow, are given in order that the reader be able to check and compare the computational results. The curve $y = \|R(t)\|_2$ and the nonoptimal upper bound $y = \kappa_2(X)\, \|r(t)\|_\infty$ are plotted in Fig. 2, and the curve $y = \|R(t)\|_2$ and the optimal upper bound $y = M^*_{2,\infty}\, \|r(t)\|_\infty = f^*_{2,\infty}\, \kappa_2(X)\, \|r(t)\|_\infty$ in Fig. 3.

(ii) *Case* $\|R(t)\|_2$ *and* $\|r(t)\|_2$: Since $\|r(t)\|_\infty \leqslant \|r(t)\|_2$, we get $\|R(t)\|_2 \leqslant \kappa_2(X)\, \|r(t)\|_2,\ t \geqslant 0$, that is, $f_{2,2} = 1$ in (64). The minimal constant $f^*_{2,2}$ in (64) and the associated time value $t^*_{2,2}$ are computed similarly as above. We obtain

$$t^*_{2,2} \doteq 0.77259828193186,$$

$$f^*_{2,2} \doteq 0.51615175566176,$$

$$M^*_{2,2} := f^*_{2,2}\, \kappa_2(X) \doteq 1.28316731275614.$$

The curve $y = \|R(t)\|_2$ and the nonoptimal upper bound $y = \kappa_2(X)\, \|r(t)\|_2$ are plotted in Fig. 4 and the curve $y = \|R(t)\|_2$ and the optimal upper bound $y = M^*_{2,2}\, \|r(t)\|_2 = f^*_{2,2}\, \kappa_2(X)\, \|r(t)\|_2$ in Fig. 5.

(iii) *Case* $\|R(t)\|_\infty$ *and* $\|r(t)\|_\infty$: For $f_{\infty,\infty} = 1.8$, in (65) we get a nonoptimal upper bound on $y = \|R(t)\|_\infty$ as can be seen from Fig. 6. The minimal constant $f^*_{\infty,\infty}$ in (65) and the associated time value $t^*_{\infty,\infty}$ are computed similarly as above. We obtain

$$t^*_{\infty,\infty} \doteq 5.49300727840584,$$

$$f^*_{\infty,\infty} \doteq 1.70885402095387,$$

$$M^*_{\infty,\infty} := f^*_{\infty,\infty}\, \kappa_2(X) \doteq 4.24825760623171.$$

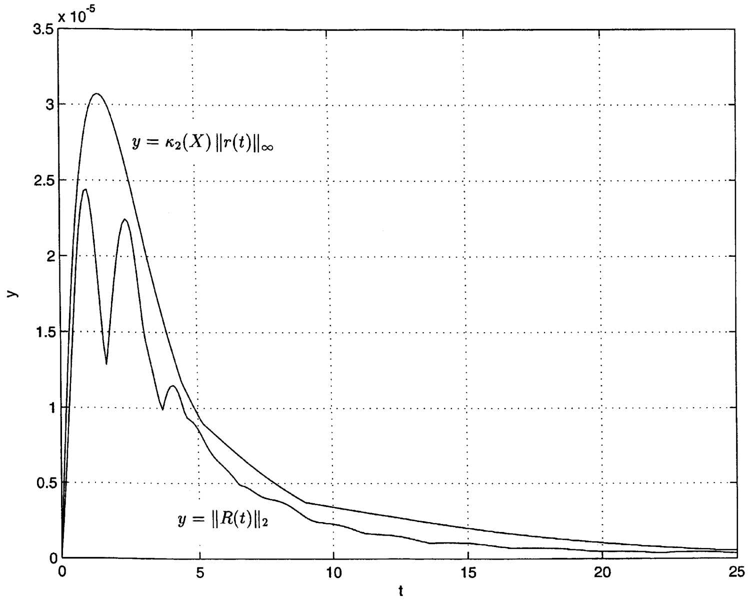

Fig. 2. $y = \|R(t)\|_2$ and *nonoptimal* upper bound $y = \kappa_2(X)\,\|r(t)\|_\infty$.

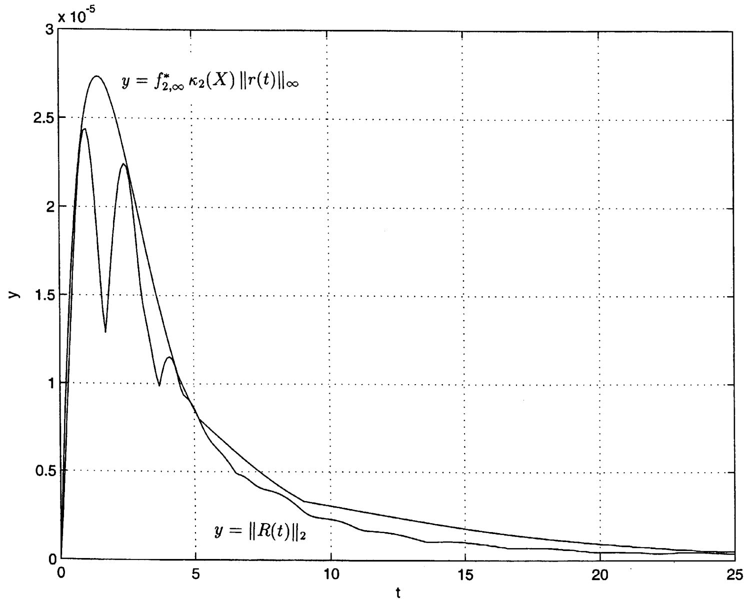

Fig. 3. $y = \|R(t)\|_2$ and *optimal* upper bound $y = f^*_{2,\infty}\,\kappa_2(X)\,\|r(t)\|_\infty$.

The curve $y = \|R(t)\|_\infty$ and the optimal upper bound $y = M^*_{\infty,\infty}\,\|r(t)\|_\infty = f^*_{\infty,\infty}\,\kappa_2(X)\,\|r(t)\|_\infty$ are plotted in Fig. 7.

(iv) *Case* $|R(t)|_2$ *and* $\|r(t)\|_2$: From Fig. 8, we see that $y = \tilde{f}_{2,2}\,\kappa_2(X)\,\|r(t)\|_2$ with $\tilde{f}_{2,2} = 1$ is a nonoptimal upper bound on $y = |R(t)|_2$. The minimal constant $\tilde{f}_{2,2} = \tilde{f}^*_{2,2}$ in (66) and the associated

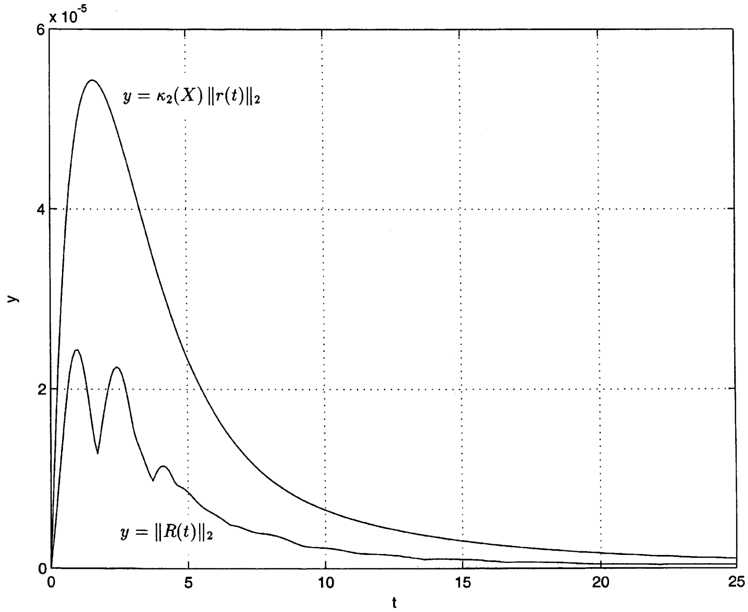

Fig. 4. $y = \|R(t)\|_2$ and *nonoptimal* upper bound $y = \kappa_2(X)\,\|r(t)\|_2$.

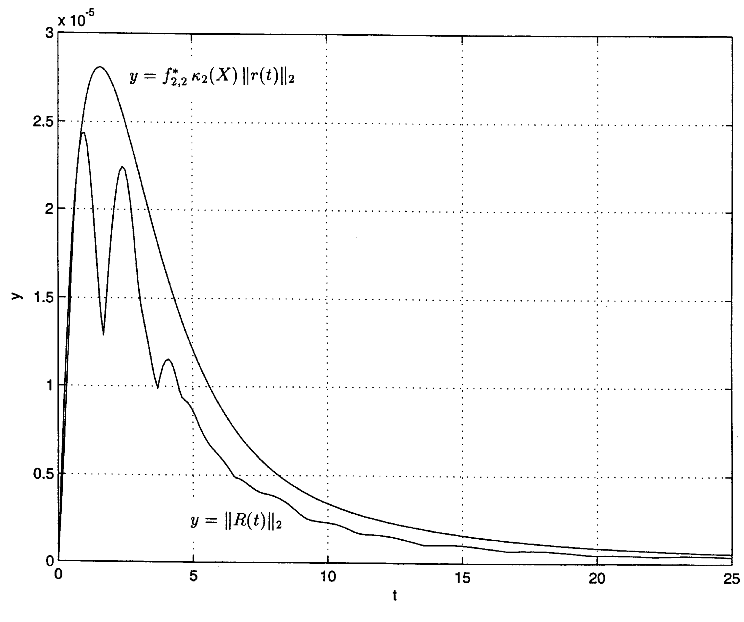

Fig. 5. $y = \|R(t)\|_2$ and *optimal* upper bound $y = f^*_{2,2}\,\kappa_2(X)\,\|r(t)\|_2$.

time value $\tilde{t}^*_{2,2}$ as well as the pertinent $\tilde{M}^*_{2,2}$-value are given by

$$\tilde{t}^*_{2,2} \doteq 0.82612246545860,$$

$$\tilde{f}^*_{2,2} \doteq 0.61247858289357,$$

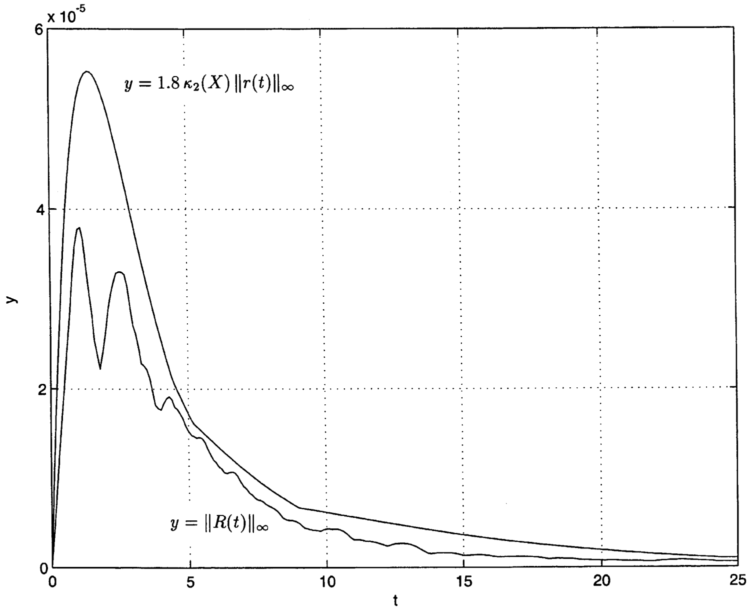

Fig. 6. $y = \|R(t)\|_\infty$ and *nonoptimal* upper bound $y = 1.8\,\kappa_2(X)\,\|r(t)\|_\infty$.

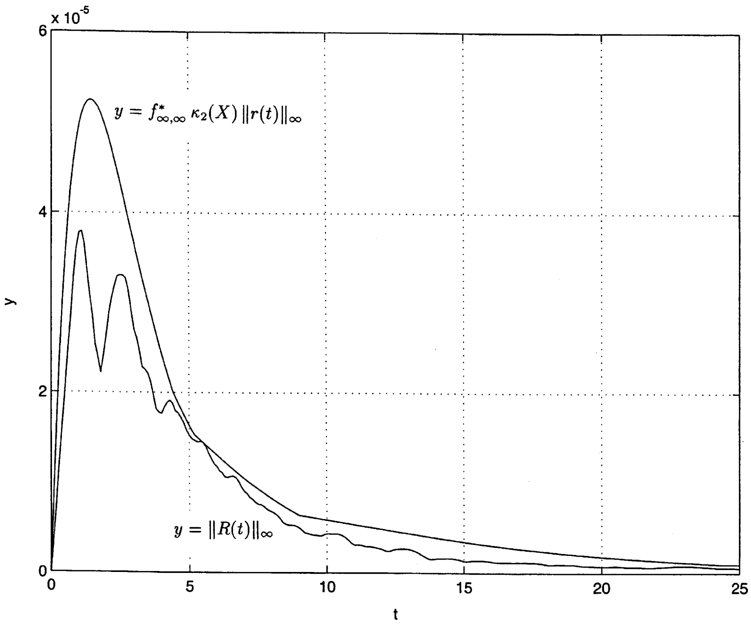

Fig. 7. $y = \|R(t)\|_\infty$ and *optimal* upper bound $y = f^*_{\infty,\infty}\,\kappa_2(X)\,\|r(t)\|_\infty$.

$$\tilde{M}^*_{2,2} := \tilde{f}^*_{2,2}\,\kappa_2(X) \doteq 1.52263842699638.$$

The curve $y = |R(t)|_2$ and the optimal upper bound $y = \tilde{M}^*_{2,2}\,\|r(t)\|_2 = \tilde{f}^*_{2,2}\,\kappa_2(X)\,\|r(t)\|_2$ are plotted in Fig. 9.

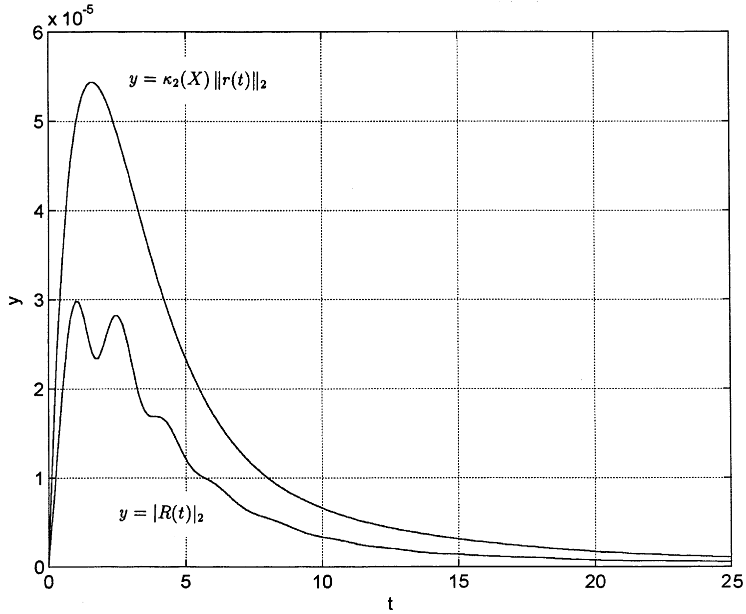

Fig. 8. $y = |R(t)|_2$ and *nonoptimal* upper bound $y = \kappa_2(X)\,\|r(t)\|_2$.

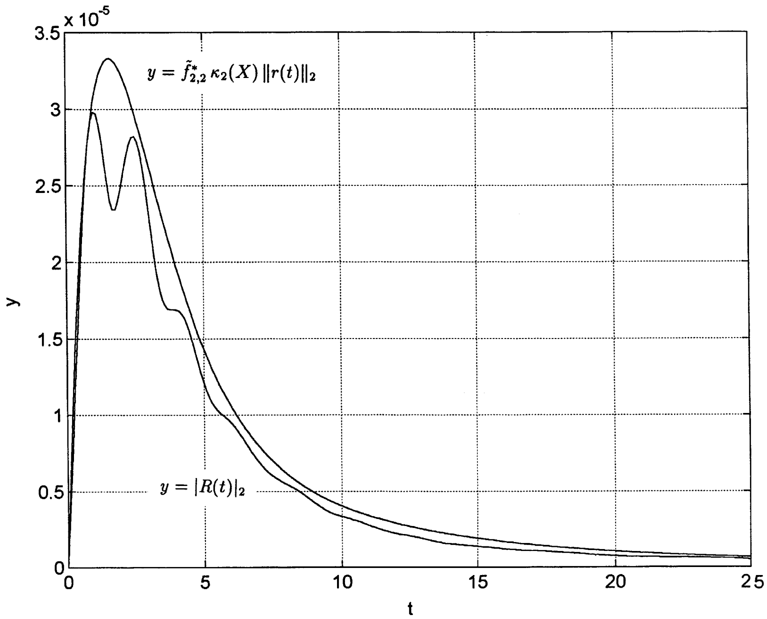

Fig. 9. $y = |R(t)|_2$ and *optimal* upper bound $y = \tilde{f}^*_{2,2}\,\kappa_2(X)\,\|r(t)\|_2$.

Table 1
Computation time for some operations

$\|R(t)\|$	$\|r(t)\|$	(t^*, f^*)	t(plot)
$\|R(t)\|_2$	$\|r(t)\|_\infty$	4.17	12.08
$\|R(t)\|_2$	$\|r(t)\|_2$	2.53	10.60
$\|R(t)\|_\infty$	$\|r(t)\|_\infty$	2.86	11.70
$\lvert R(t)\rvert_2$	$\|r(t)\|_2$	2.09	9.34

Table 2
Comparison between $(\Psi_{\Delta t})^\delta$ and $\Psi_{\delta\Delta t}$

δ	0.5	t
$\|(\Psi_{\Delta t})^\delta\|_2$	1.32037295919365	0.05
$\|\Psi_{\delta\Delta t}\|_2$	1.32035774554620	0.05
$D_+\|(\Psi_{\Delta t})^\delta\|_2$	0.13052154263271	0.11
$D_+\|\Psi_{\delta\Delta t}\|_2$	0.13049219502112	0.11

On the whole, one can say that the nonoptimal upper bounds can be improved in all cases such that the best possible upper bounds are obtained.

5.4. Computational aspects

In this subsection, we say something about the used computer equipment, the computation time for some operations and on the approximation of noninteger powers of the discrete evolution.

(i) As to the *computer equipment*, the following hardware was available: a Pentium II CPU at 300 MHz, an 8 GB mass storage facility, two SDRAM 64 MB high-speed memory. As software package, we used 368-Matlab, Version 4.2c.

(ii) The *computation time* t of an operation was determined by the command sequence $t0 = clock$; *operation*, $t = etime(clock, t0)$; it is put out in seconds rounded to two decimal places, by MATLAB. Let $t(t^*, f^*)$ be the computation time for the determination of (t^*, f^*) (e.g., $(t^*, f^*) = (t^*_{2,\infty}, f^*_{2,\infty})$ in case (i) of Section 5.3), and let t(plot) be the computation time for determining the (t, y)-values for the plot of one of the Figs. 2–9 which includes the time $t(t^*, f^*)$. Then, for the four cases in Section 5.3, we obtained the results in Table 1.

(iii) The referee was kind enough to point out that for $0 < \delta < 1$ the *noninteger power of the discrete evolution* $(\Psi_{\Delta t})^\delta$ can be avoided by the *approximation* $\Psi_{\delta\Delta t}$. In this respect, we have compared the computational results and the computation times t. For $\Delta t = 0.1$ and $\delta = 0.5$, we obtained the results in Table 2.

So, there was no measurable difference, i.e., the time difference was less than 1/100 second. Further, if $(\Psi_{\Delta t})^\delta$ is replaced by $\Psi_{\delta\Delta t}$, the result is less precise. As a consequence, at least for our purpose, there was no reason to approximate $(\Psi_{\Delta t})^\delta$ by $\Psi_{\delta\,\Delta t}$. However, the proposed approximation could be of interest in other problems.

Acknowledgements

The author would like to give thanks to the referee for his/her detailed comments on a former draft of this paper, which led to better presentation.

References

[1] W.A. Coppel, Stability and Asymptotic Behavior of Differential Equations, D.C. Heath, Boston, 1965.
[2] G. Dahlquist, Stability and Error Bounds in the Numerical Integration of Ordinary Differential Equations, Kungl. Tekn. Högsk. Handl., Trans. Royal Inst. Technol. No. 130, Stockholm, 1959.
[3] Ch.A. Desoer, H. Haneda, The measure of a matrix as a tool to analyse computer algorithms for circuit analysis, IEEE Trans. Circuit Theory 19 (5) (1972) 480–486.
[4] P. Deuflhard, F. Bornemann, Numerische Mathematik II, Walter de Gruyter, Berlin, New York, 1994.
[5] G.H. Golub, Ch.F. van Loan, Matrix Computations, The Johns Hopkins University Press, Baltimore, MD, 1989.
[6] E. Hairer, S.P. Nørset, G. Wanner, Solving Ordinary Differential Equations I, Springer, Berlin, 1993.
[7] I. Higueras, B. García-Celayeta, Logarithmic norms for matrix pencils, SIAM J. Matrix Anal. Appl. 20 (1999) 646–666.
[8] I. Higueras, B. García-Celayeta, How close can the logarithmic norm of a matrix pencil come to the spectral abscissa? SIAM J. Matrix Anal. Appl. 22 (2000) 472–478.
[9] Guang-Da Hu, Guang-Di Hu, A relation between the weighted logarithmic norm of a matrix and the Lyapunov equation, BIT 40 (3) (2000) 606–610.
[10] T. Kato, Perturbation Theory for Linear Operators, Springer, New York, 1966.
[11] L. Kohaupt, Second logarithmic derivative of a complex matrix in the Chebyshev norm, SIAM J. Matrix Anal. Appl. 21 (2) (1999) 382–389.
[12] L. Kohaupt, Differential calculus for some p-norms of the fundamental matrix with applications, J. Comput. Appl. Math. 135 (2001) 1–21.
[13] L. Kohaupt, Differential calculus for p-norms of complex-valued vector functions with applications, J. Comput. Appl. Math. 145 (2002) 425–457.
[14] S.M. Lozinskiĭ, Error estimates for the numerical integration of ordinary differential equations (Russian), I. Izv. Vysš. Učebn. Zaved. Mat. 5 (6) (1958) 52–90.
[15] P.C. Müller, W.O. Schiehlen, Lineare Schwingungen, Akademische Verlagsgesellschaft, Wiesbaden, 1976.
[16] H. Niemeyer, E. Wermuth, Lineare Algebra, Vieweg, Braunschweig, Wiesbaden, 1987.
[17] C.V. Pao, Logarithmic derivatives of a square matrix, Linear Algebra Appl. 6 (1973) 159–164.
[18] C.V. Pao, A further remark on the logarithmic derivatives of a square matrix, Linear Algebra Appl. 7 (1973) 275–278.
[19] J. Stoer, Einführung in die Numerische Mathematik, Vol. I, Springer, Berlin, Heidelberg, New York, 1972.
[20] T. Ström, Minimization of norms and logarithmic norms by diagonal similarities, Computing 10 (1972) 1–7.
[21] T. Ström, On logarithmic norms, SIAM J. Numer. Anal. 10 (5) (1975) 741–753.
[22] A.E. Taylor, Introduction to Functional Analysis, Wiley, New York, London, 1958.
[23] H. Waller, W. Krings, Matrizenmethoden in der Maschinen- und Bauwerksdynamik, Bibliographisches Institut, Mannheim, Wien, Zürich, 1975.
[24] J.H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1965.